[153] A. Nicholson, Q. Stiévenart, A. Mazidi, and M. Ghafari, "Wasmizer: Curating WebAssembly-driven Projects on GitHub," in *2023 IEEE/ACM 20th International Conference on Mining Software Repositories (MSR)*, pp. 130–141, 2023.

[154] T. Y. Zhuo, Z. Yang, Z. Sun, Y. Wang, L. Li, X. Du, Z. Xing, and D. Lo, "Source Code Data Augmentation for Deep Learning: A Survey," *arXiv e-prints*, p. arXiv:2305.19915, May 2023.

[155] S. Srikant, S. Liu, T. Mitrovska, S. Chang, Q. Fan, G. Zhang, and U. O'Reilly, "Generating Adversarial Computer Programs using Optimized Obfuscations," in *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*, OpenReview.net, 2021.

[156] H. Ye, M. Martinez, X. Luo, T. Zhang, and M. Monperrus, "SelfAPR: Self-supervised Program Repair with Test Execution Diagnostics," in *37th IEEE/ACM International Conference on Automated Software Engineering, ASE 2022, Rochester, MI, USA, October 10-14, 2022*, pp. 92:1–92:13, ACM, 2022.

[157] W. Zhang, S. Guo, H. Zhang, Y. Sui, Y. Xue, and Y. Xu, "Challenging Machine Learning-based Clone Detectors via Semantic-preserving Code Transformations," *IEEE Trans. Software Eng.*, vol. 49, no. 5, pp. 3052–3070, 2023.

[158] H. Li, X. Zhou, L. A. Tuan, and C. Miao, "Rethinking Negative Pairs in Code Search," *arXiv preprint arXiv:2310.08069*, 2023.

[159] J. D. Seideman, *Transformation and Abstraction to Aid Comparison of Binary Executables Across Compilation Environments.* PhD thesis, City University of New York, 2023.

[160] H. Huang, A. M. Youssef, and M. Debbabi, "BinSequence: Fast, Accurate and Scalable Binary Code Reuse Detection," *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, 2017.

[161] J. Jang, A. Agrawal, and D. Brumley, "ReDeBug: Finding Unpatched Code Clones in Entire OS Distributions," in *2012 IEEE Symposium on Security and Privacy*, pp. 48–62, 2012.

[162] H. Jang, K. Yang, G. Lee, Y. Na, J. D. Seideman, S. Luo, H. Lee, and S. Dietrich, "QuickBCC: Quick and Scalable Binary Vulnerable Code Clone Detection," in *ICT Systems Security and Privacy Protection*, pp. 66–82, 2021.

Part II

Included papers